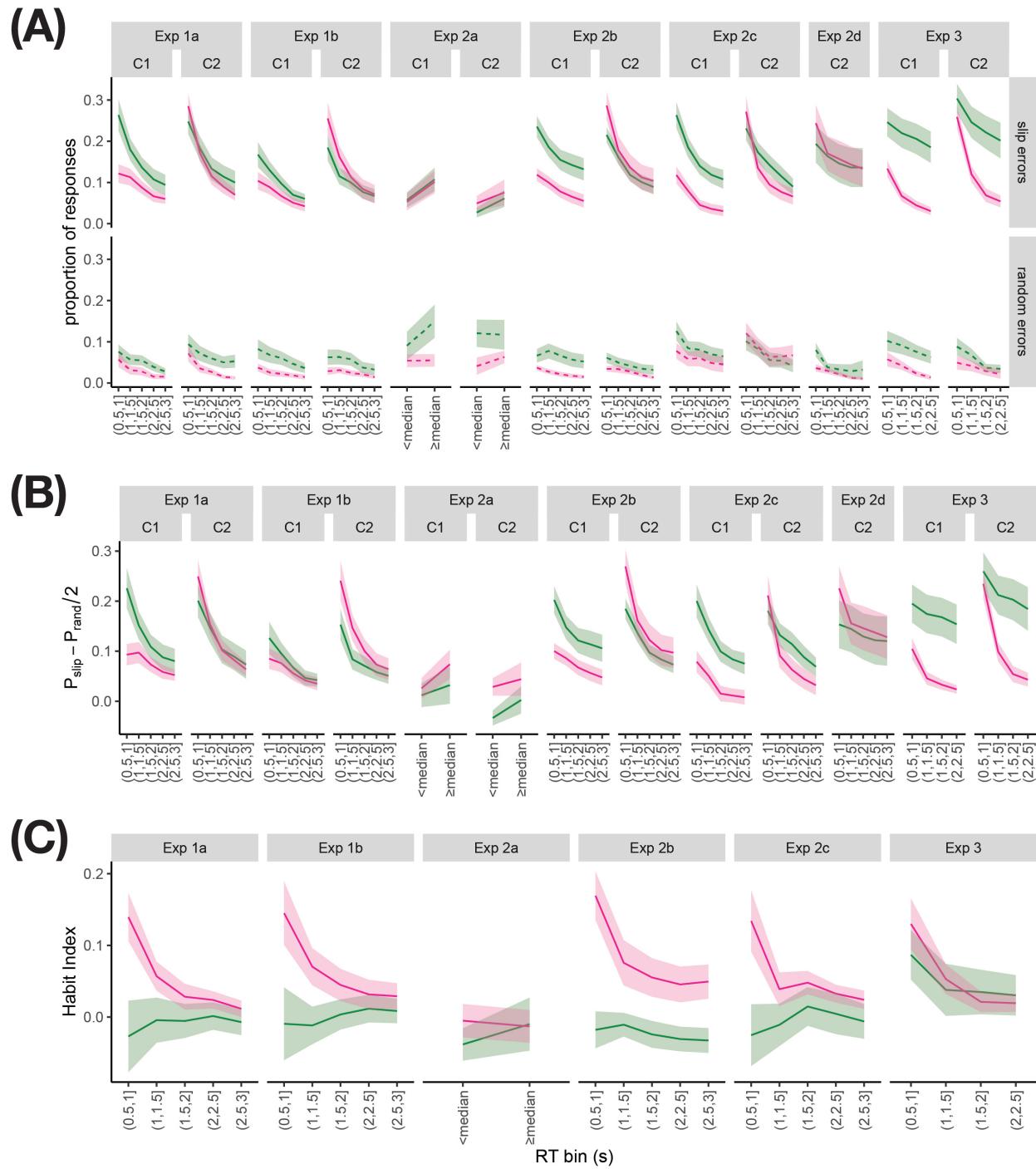
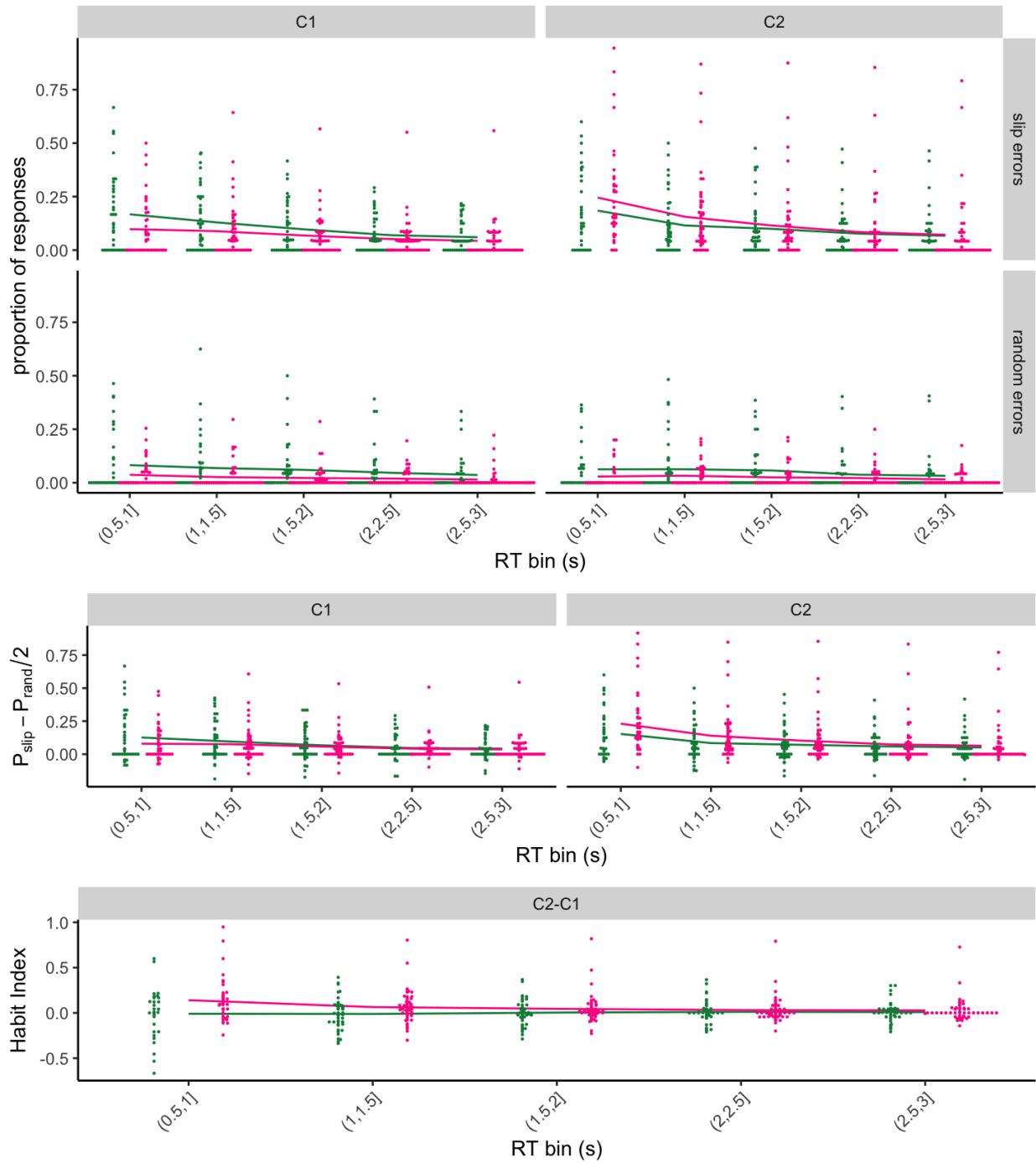


1190

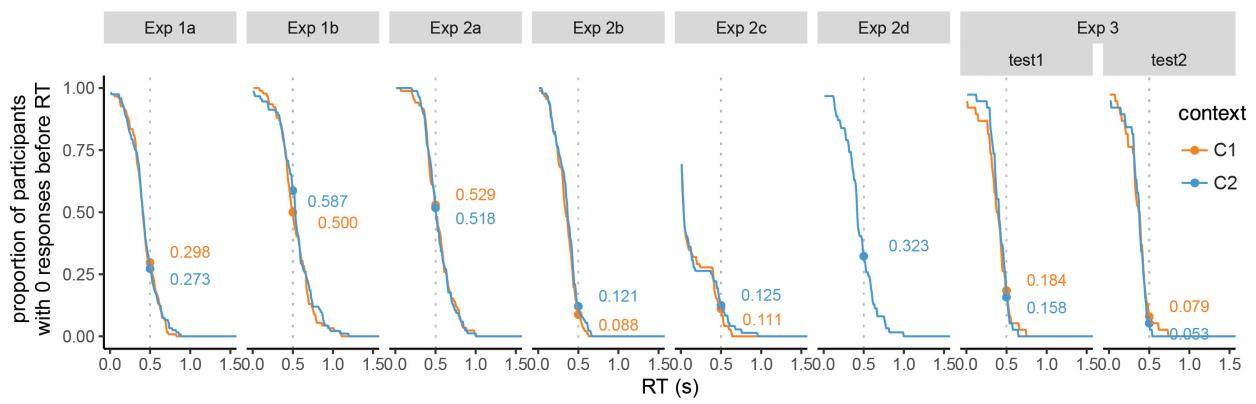
## Supplementary Information

**Figure S1**

*Habit Index and quantities used to compute the Habit Index, across task versions (mean  $\pm$  s.e.m.). For full legend, see Fig. 3*

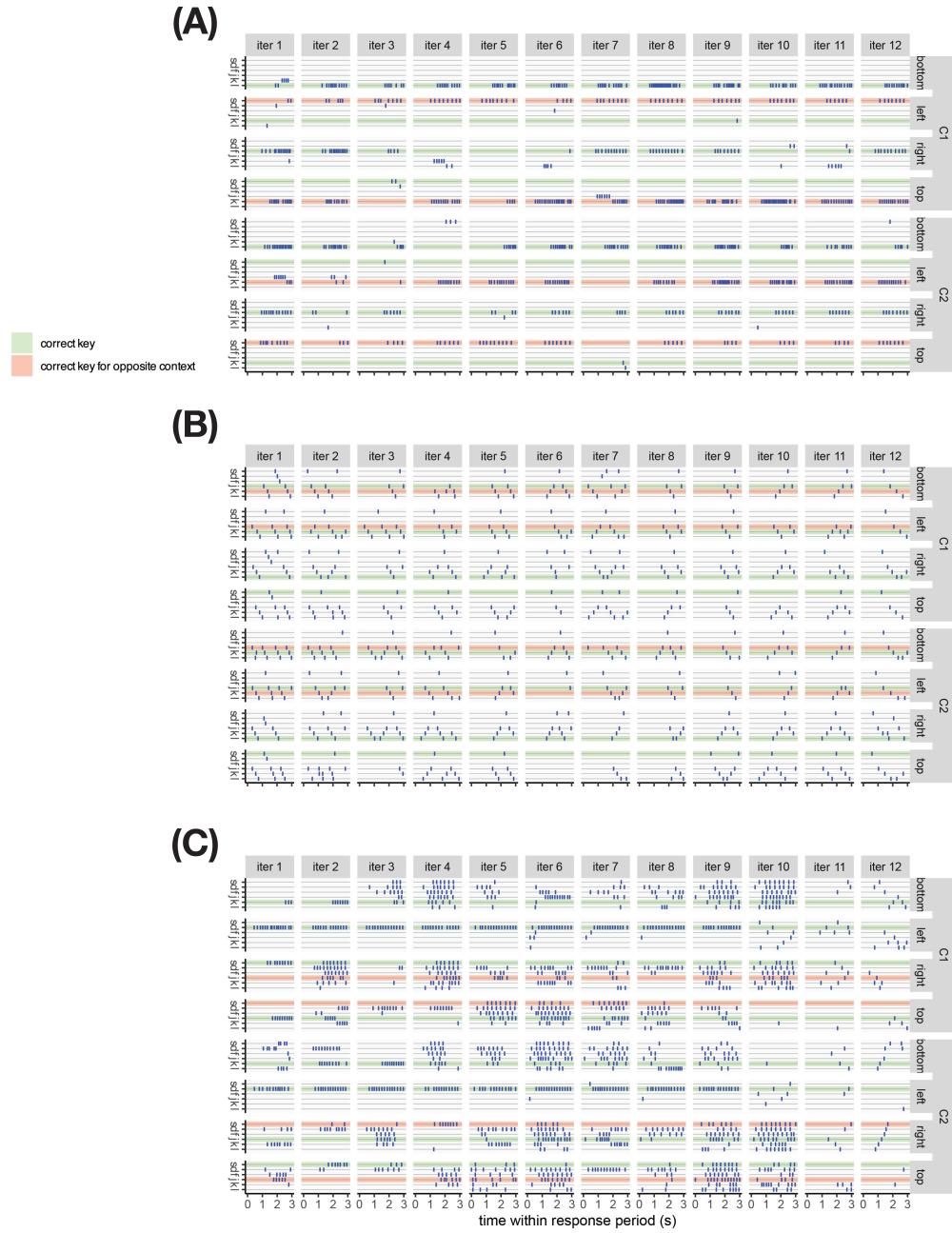
**Figure S2**

*Data for Fig. 3, with individual data points for Exp 1b, showing non-normal variability across participants.*



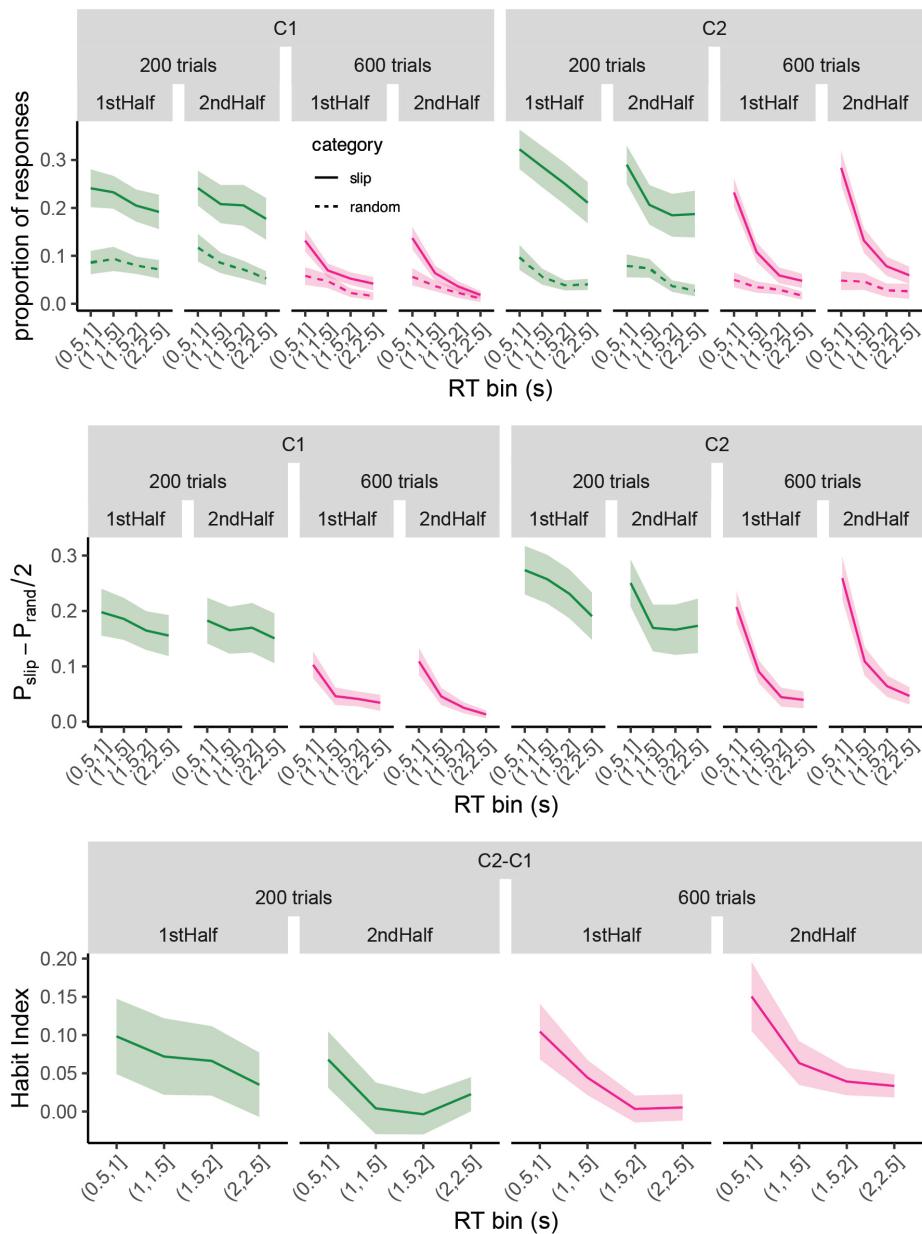
**Figure S3**

*Proportion of participants with 0 responses before RTs spanning the response period, for each Experiment. We used this analysis to select the “early” time bin of [.5, 1], as we found that a large proportion of participants did not have enough data with RTs faster than 0.5s.*



**Figure S4**

All test phase stage 3 responses from example participants excluded for satisficing behaviors (see Exclusion Criteria in main text for a detailed description of each). Incongruent arms have both a green row (indicating correct response for current context) and a red row (indicating correct response for opposite context); congruent arms have only a green row (the correct response is the same across contexts). **(A)** consistent errors: the participant ended on the incorrect response on all trials for both incongruent arms (left and top) in both contexts. **(B)** random spam: the participant tried pressing multiple keys regardless of congruent/incongruent arm type, allowing them to collect points without making an effort to press the correct key. **(C)** strategic spam: the participant was classified as a strategic spammer because they tried pressing multiple keys more often than not when in incongruent arms, but not in congruent arms.

**Figure S5**

Error rates, baseline-corrected slip rate, and Habit Index in Exp 3, broken down by subblocks each covering half of one of the two test phases.

<sub>1191</sub> A short (approximately 5 minutes long) demo of Experiment 1b can be found at  
<sub>1192</sub> [sarahoh.github.io/habitTask\\_demo/exp.html](https://sarahoh.github.io/habitTask_demo/exp.html).